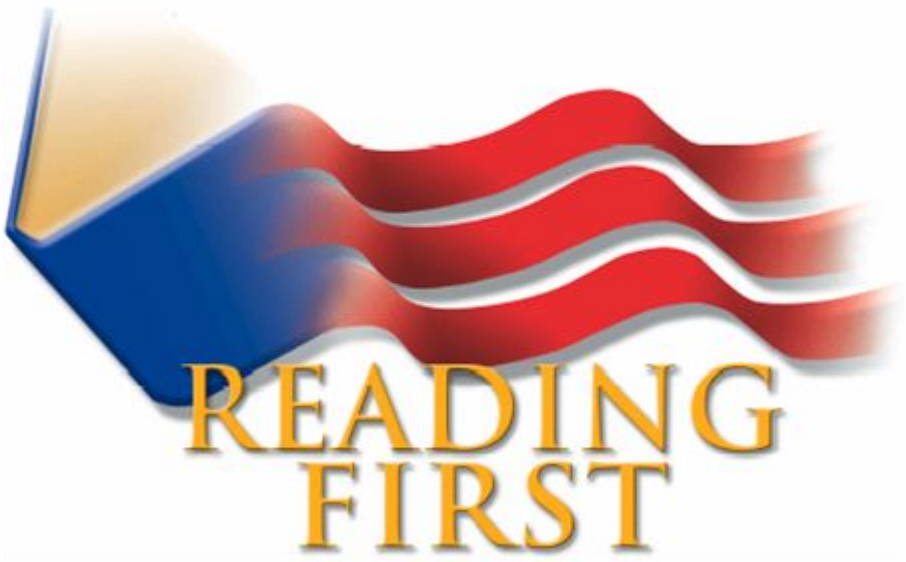
The Iowa Department of Education's Reading First State Evaluation



The Psychology in Education Research Lab (PERL) serves as the external evaluator for the State of Iowa's Reading First Program. PERL is a research laboratory within the Department of Curriculum & Instruction, College of Human Sciences at Iowa State University. Personnel at PERL are comprised of researchers who are knowledgeable in the fields of education and psychology. Psychological research methods and design are applied in educational settings in examining the impact of school interventions in increasing student achievement in reading. PERL researchers have extensive experience in designing, implementing, and evaluating scientifically based research. The following is a synopsis of the evaluation component of the State of Iowa's Reading First Program.

Reading First Data Collection Plan

The activities a building engages in revolve around its area of instructional focus, the study of teacher implementation of a targeted strategy(ies) and its effects on student achievement. The Every Child Reads: Excellence in Teaching and Learning Data Collection Plan (See Figure 1) was designed for two purposes:

1. To provide a minimum target for schools when studying student and staff performance within structured school improvement efforts
2. To determine the level of implementation of a school's efforts within a focus area in order to provide ongoing support

The timeline of the data collection plan is directly aligned with the student achievement data collected for Reading First Schools in the fall and spring.

Web-based Data Collection Center

Central to the evaluation of the Iowa Reading First Program is the collection of student data. PERL has developed, implemented and maintained a secure (password-protected) web-based data collection center. To assist schools to navigate through the web site (e.g., data entry, running reports, charts), PERL provides support and a user manual that is easy to follow. Training is provided as needed.

Student data is collected two times per year (fall and spring) aligned with the Data Collection Plan. Tests administered include the Phonological Awareness Test (PAT), Basic Reading Inventory (BRI), and Iowa Tests of Basic Skills (ITBS). Appendix A describes the tests in further detail. Figure 2 indicates the tests that are administered in the fall and spring by grade.

Sociodemographic data is also collected on each student. Sociodemographic data collected includes gender, students with/without disabilities, major race/ethnic categories, students with economic advantage/disadvantage, and students with/without English limited proficiency. In addition, specific information regarding special education status, referral for pre-referral services is also collected. The sociodemographic data allows us to disaggregate tests scores by these five subgroups.

Web-based Reports

PERL has provided schools and districts with the ability to generate building/district level reports. Report options include the number and percentage of students at grade level (agl), in need of additional intervention (nai), and need substantial intervention (nsi) by test and by grade.

Results can be disaggregated by the five categories (i.e., gender, economic advantage/disadvantage, students with/without disabilities, student with/without limited English proficiency, major race/ethnic categories) identified in the federal Reading First funding requirements.

Buildings/districts also have chart options that include percentage of students proficient by test, trend lines of the percentage of students by time, percentage of students proficient by the disaggregated groups, and the percentage of students at or below proficiency by time. Both reports and charts can be generated and dropped into a manuscript or Word document.

Student Data Analysis

On a yearly basis, the test data and demographic data are analyzed to determine progress made by schools to increase the percentage of students proficient in reading as well as decreasing the achievement gap between groups (e.g., students with disabilities versus students without disabilities).

Schools are evaluated to determine whether they were able to meet performance benchmarks on each test (by grade). Schools can meet performance benchmarks in one of two ways. The first method involves a statistical comparison of the percentage of students proficient in the fall to the percentage of students proficient in the spring. The second method involves determining whether 75% or more of the students were proficient in the spring. (For more information on Performance Benchmarks, please see Appendix B.)

Analysis is also conducted to determine which schools made the greatest gains in increasing the percentage of students proficient in each of the tests between Fall and Spring or between 2003-2004 (Year1) and 2004-2005 (Year2).

Study of Implementation

Currently, PERL is evaluating the impact of technical assistance/teacher implementation in increasing student achievement in reading. The data used in these analyses derive from monthly Technical Assistance Reports submitted electronically by each Reading First School and building profiles indicating the extent or level that data was used to guide actions in buildings, compiled by AEA reading teams during the academic year (See Appendices C & D). Analyses of these data will include both a cross-sectional (i.e., one point in time) and longitudinal (i.e., change over time) examination of the impact.

Annual Report

An important role of the external evaluator is to assist with the preparation of the annual report required by the funding agency. The external evaluator provides an executive summary addressing the overall effectiveness of the Iowa Reading First program. In addition, PERL prepares and submits a data set containing the number of students proficient by test, grade, and school along with the federal annual report.

Collaboration with Iowa Department of Education

The relationship between the PERL and the Iowa Department of Education is collaborative in nature. Collaboratively, the Iowa Department of Education and PERL make recommendations and decisions in terms of collecting student achievement data, measurement, defining greatest gains, and performance benchmarks as well as review the strengths of the program and appropriateness of testing.

Figure 1.

Every Child Reads: Excellence In Teaching and Learning

2005-2006 Data Collection Plan

In order to improve student achievement, school staff must study both teacher implementation within a focus area and its effects on student achievement. This table is designed for two purposes:

1. To provide a minimum target for a school when studying student and staff performance within structured school improvement efforts.
2. To determine the level of implementation of a school's efforts within a focus area in order to provide ongoing support.

Time Frame	Data
August 2, 2005– November 11, 2005	▪ Data Tables 1–9, and 15 *(Completion of the Analyzing and Reporting Our Data Structured Response Sheet is optional.)* ▪ Diagnostic Assessment Data* with a completed Analyzing and Reporting Our Data Structured Response Sheet ▪ Classroom Collection Data (Access to Print) with a completed Analyzing and Reporting Our Data Structured Response Sheet ▪ Initial record of your 2005–2006 building profile and review of your 2005-2006 action plans
November 14, 2005– February 24, 2006	▪ First study of Teacher Implementation within focus area with a completed Analyzing and Reporting Our Data Structured Response Sheet ▪ Second study of Teacher Implementation within focus area with a completed Analyzing and Reporting Our Data Structured Response Sheet ▪ *Optional: Diagnostic Assessment Data and completion of the Analyzing and Reporting Our Data Structured Response Sheet* ▪ NRT Data with a completed Analyzing and Reporting Our Data Structured Response Sheet if administered in the fall (Tables 21, 22, 22b, 29, & 30)† ▪ Update your 2005-2006 building profile and review of your 2005–2006 action plans
February 27, 2006– April 28, 2006	▪ Third study of Teacher Implementation within focus area with a completed Analyzing and Reporting Our Data Structured Response Sheet ▪ Diagnostic Assessment Data** with a completed Analyzing and Reporting Our Data Structured Response Sheet ▪ Update your 2005-2006 building profile and review of your 2005–2006 action plans
End of the year	▪ Fourth study of Teacher Implementation within focus area with a completed Analyzing and Reporting Our Data Structured Response Sheet ▪ 2006-2007 Action Plans with Action Plan Analysis & Rating Scale ▪ NRT Data with a completed Analyzing and Reporting Our Data Structured Response Sheet if administered in the spring (Tables 21, 22, 22b, 29, & 30) † ▪ Update your 2005–2006 building profile

Notes:

**Diagnostic Data Requirements: Fall*
1) Phonemic Awareness Data: Grades K–1
2) Fluency Data: Grades 2–3
3) Comprehension Data: Grades 2–3

***Diagnostic Data Requirements: Spring*
1) Phonemic Awareness Data: Grades K–1
2) Fluency Data: Grades 1, 2, 3
3) Comprehension Data: Grades 1, 2, 3

† *Analysis for ITBS should look across years (2003/2004 to 2004/2005 to 2005/2006) and for this year (2005/2006).*

Figure 2.

Assessments By Grade Level Reading First State Evaluation Schedule

The following table indicates the tests required in Fall and Spring by grade for Reading First State Evaluation purposes.

TEST	FALL					SPRING				
	K	1	2	3	4	K	1	2	3	4
Phonological Awareness										
Rhyming	X	X				X	X			
Deletion	X	X				X	X			
Blending	X	X				X	X			
Segmentation		X					X			
Isolation		X					X			
Substitution		X					X			
Phonics										
Graphemes		X	X				X			
Decoding		X	X				X			
BRI										
Fluency (Grade level passage)			X	X			X	X	X	
Comprehension (Grade level passage)			X	X			X	X	X	
ITBS										
Reading Total (NPR & IPR)				**See Note**					**See Note**	
Reading Comprehension (NPR & IPR)										
Vocabulary (NPR & IPR)										

Note: ITBS is required for 3rd and 4th graders, however it is only administered once per year. Schools determine when the ITBS is administered.

Appendix A

Phonological Awareness Test (Phonological Awareness and Phonics)

The Phonological Awareness Test is a normed referenced test designed to assess phonological processing and phoneme-grapheme correspondence (Robertson & Salter, 1997).

The following phonological processing subtests are administered to kindergarten and first grade students: rhyming, deletion, and blending. Some of the phonological processing subtests may not be appropriate for all five year olds; therefore, the following subtests are only administered to first graders: segmentation, isolation, and substitution.

The phonics subtests (graphemes and decoding) are administered to first graders in the fall and spring. For scoring purposes, students who are proficient in phonological processing and phoneme-grapheme correspondence are not re-tested during subsequent testing. A presumption is made that students whose scores indicate they are proficient in a particular subtest have mastered this skill and no longer require testing. Therefore, the number of students who pass in the fall are added to the number of student who passed in the spring.

Basic Reading Inventory (Reading Fluency and Comprehension)

To assess student achievement in reading fluency and comprehension, the Basic Reading Inventory (BRI) was administered to second and third graders in the fall and spring. The BRI is an informal reading assessment test comprised of a series of graded word lists and graded passages that can be used to gain insight into these areas (Johns, 2001). Student scores reported reflect whether students were independent at their current grade level in fluency and comprehension.

Iowa Tests of Basic Skills (ITBS)

The Iowa Tests of Basic Skills is an achievement battery of tests comprised of various subject areas that have been all been standardized within the same group of students (Hoover, H., Dunbar, S., Frisbie, D., Oberley, K., Bray, R., Naylor, J., Lewis, J., Ordman, V., & Qualls, A.L., 2003).

National and Iowa percentile rank scores are derived from each of the following reading subject areas: vocabulary, reading, and comprehension. The vocabulary test is a measure of a students' reading vocabulary. The Reading Total subtest assesses the extent of student's development in reading comprehension. The comprehension test assesses three main skills: Factual Understanding, Inference and Interpretation, and Analysis and Generalization.

Students in the third and fourth grades were administered the ITBS once during the fall, winter, or spring of the 2003-2004 school year. Districts/schools determine the time of the year it is administered in their respective districts/schools.

References

Hoover, H., Dunbar, S., Frisbie, D., Oberley, K., Bray, R., Naylor, J., Lewis, J., Ordman, V., & Qualls, A.L., (2003). *The Iowa Tests: Interpretive Guide for School Administrators.* Riverside, CA: Riverside Publishing.

Johns, J. (2001). *Basic Reading Inventory: Pre-primer through grade twelve and early literacy assessments*. Dubuque, IA: Kendall/Hunt Publishing Company.

Robertson, C. & Salter, W. (1997). *The Phonological Awareness Test.* East Moline, IL: LinguiSystems, Inc.

Appendix B

Understanding Performance Benchmarks and Their Use for Reading First Schools

What is the purpose for performance benchmarks?

For Reading First Schools, performance benchmarking is used to determine if there is a statistically significant increase in the proportion of students attaining proficiency and to determine a building's funding status.

How do we determine whether performance benchmarks have been met?

Schools can meet their performance benchmarks in one of two ways. The first method involves a statistical comparison of the percentage of students proficient over time. The second method involves determining whether 75% or more of the students were proficient on the identified assessment.

The percentage of students proficient in Time 1 is statistically compared to the percentage of students proficient in Time 2. Schools that achieve a statistically significant increase between fall and spring are coded as having met their performance benchmark. Comparisons are made by test and by grade.

When schools do not meet their performance benchmarks statistically, the second method of assessment is used. Schools with 75% or more of their students proficient at Time 2 are coded as having met their performance benchmark. These analyses are made by test and by grade. The second method is used because some schools will not be able to statistically increase the percentage of students proficient from Time 1 to Time 2. In particular, school size and the percentage of students proficient at baseline may affect whether schools are able to increase the percentage of students proficient at Time 2 statistically. Sample size affects significance testing and smaller schools may have greater difficulty meeting their performance benchmark statistically (see "Sample size influences whether statistically significant differences are achieved"). Other schools will not be able to significantly increase the percentage of students proficient at Time 2 because they have a relatively large percentage of students who are proficient on their tests at baseline (e.g., fall). As a result these schools will make smaller gains at Time 2 making it impossible to achieve a statistically significant difference. However, the percentage of students proficient at these schools may be greater than the percentage of students proficient among some of the schools that achieved statistical significance.

What does statistically significant difference mean?

Statistical significance is important when comparisons between groups (i.e., the percentage of student proficient in the fall compared to the percentage proficient in the spring) are made. When statistically significant differences between groups are found we can be confident that the difference did not occur as a result of chance. In other words, we would expect similar results in subsequent testing among these students.

Sample size influences whether statistically significant differences are achieved.

One of the biggest challenges in determining whether there is significant difference between Time 1 and Time 2 involves the size of the sample (i.e., number of students) used in the analysis. In general, the larger the sample size, the higher the probability that statistically significant differences can be found. Schools with smaller sample sizes have to achieve greater differences (i.e., move more students) between the percentage of students proficient in the fall compared to the percentage of students proficient in the spring in order for the difference to be statistically significant. For example, similar proportions of students proficient between fall (50%) and spring (71%) may be significantly different at a larger school (n = 60) and not significantly different at a smaller school (n = 20)*.

*Note: values provided in last paragraph are for illustrative purposes only.

Appendix C

Technical Assistance Report

Month:

School:

SWRT Member's Name:

Purposes for Technical Assistance Reports

1. To gather data about the level of technical assistance provided to schools.
2. To gather data about the level of implementation at the school.

Remember to review the school folder before each visit. If something is missing (such as a school action plan or the current set of tables on general school data), do your best to collect it. This may include helping the principal of Leadership Team do the work.

Record of Contacts

Date	SWRT Members Present

Leadership Team Meetings

Date	Principal Attended	Meeting Length	Unit / Module	Activities Addressed

Staff Development Meetings

Date	Principal Attended	Meeting Length	Unit / Module	Activities Addressed

Comments

Briefly describe the staff's actions/activities:

A) Current Actions

B) Future Actions

General comments on the staff's progress:

Staff Concerns:

Actions to be taken by SWRT member based on identified needs:

Follow-up requiring a response from DE K-3 Reading Team member:

Classrooms Visited: Lessons Observed: Demonstrations Performed:

ON-SITE VISIT

District/School:

Date:

Length of Visit:

SWRT member(s) present:

Respond to the following:	**Response**	**If yes, for how long? (In minutes)**
Did you meet with the principal?		
Did you meet with the leadership team?		
Did you meet with the staff as a whole?		
What Unit/Module did you engage in?		
Activities addressed during meeting?		
Analysis of Evaluation		
Total Minutes		

Classrooms Visited: Lessons Observed: Demonstrations Performed:

Staff Member Name	**Grade Level**	**Length of Observation**	**Stategy Observed**	**Comments**

Audience	**Strategy**	**Who did it?**

Appendix D

Every Child Reads: Excellence In Teaching and Learning

Building Profile

Study of the Continuous Assessment of Student Performance & Implementation

District Name: _____ School Name:_____ Year:_____

Rating Scale:
3 points – Data present with a clear indicator of data analysis and use
2 points – Data present with a clear indicator of analysis
1 points – Data present but no indicator of analysis
0 points – No data present

Data Collected	Points Assigned	Comments
Diagnostic Data: Fall *		
Diagnostic Data: Spring**		
Student and Staff Population Data:		
Tables 1, 2, 3, 4, 5, 6, 7, 8, 9, & 15		
NRT Data: Table 21†		
NRT Data: Table 22†		
NRT Data: Table 22b†		
NRT Data: Tables 29 & 30†		
Classroom Collection Data		
First Study of Teacher Implementation		
Second Study of Teacher Implementation		
Third Study of Teacher Implementation		
Fourth Study of Teacher Implementation		
Total Points		

High Implementation: 22–34 points with a 2 or 3 on at least 11 variables
Moderate Implementation: 11–21 points with a 2 or 3 on at least 7 variables
Low Implementation: 10 points or below

Notes:

**Diagnostic Data Requirements: Fall*
1) Phonemic Awareness Data: Grades K–1
2) Fluency Data: Grades 2–3
3) Comprehension Data: Grades 2–3

***Diagnostic Data Requirements: Spring*
1) Phonemic Awareness Data: Grades K–1
2) Fluency Data: Grades 1, 2, 3
3) Comprehension Data: Grades 1, 2, 3

† *Analysis for ITBS should look across years (2003/2004 to 2004/2005 to 2005/2006) and for this year (2005/2006).*